

# DATA PROCESSING SYSTEM AND PARALLEL COMPUTER

Publication number: JP8108443 (A)

Publication date: 1996-04-23

Inventor(s): TAMURA YASUSHI; MATSUOKA HITOSHI +

Applicant(s): HITACHI LTD +

Classification:

- international: G06F13/12; G06F15/16; G06F15/163; G06F13/12; G06F15/16; (IPC1-7): G06F13/12; G06F15/163

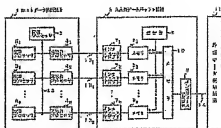
- European:

Application number: JP19940241097 19941005

Priority number(s): JP19940241097 19941005

Abstract of JP 8108443 (A)

PURPOSE: To improve the data transfer speed and the difference in speed between a host data processor which has a slow input/output transfer speed and an external data storage device which has a fast in speed. CONSTITUTION: A host data processor 1 is provided with a managing processor 2, plural arithmetic processors 31-3m, and plural input/output processors 41-4n. An input/output data buffer device 5 is equipped with a control part 6, interface adapters 71-7n and 8, plural memories 91-9n, and a memory 10. The managing processor 2 receives a READ instruction or WRITE instruction from an arithmetic processor and selects plural input/output processors required for data input/output operation. Data transferred between the host data processor and external data storage device 11 are divided into plural parts, which are transferred in parallel between the selected input/output processors of the host data processor 1 and plural memories of the input/output data buffer device 5.



Data supplied from the **espacenet** database — Worldwide

特開平8-106443

(43) 公開日 平成8年(1996)4月23日

(51) Int. Cl. <sup>6</sup> G 0 6 F 15/163 13/12	識別記号 3 4 0 B 7368-5E	F I	技術表示箇所 G 0 6 F 15/ 16 3 1 0 V
-------------------------------------------------------	-------------------------	-----	----------------------------------

審査請求 未請求 請求項の数3 O L (全 6 頁)

(21) 出願番号 特願平6-241097

(22) 出願日 平成6年(1994)10月5日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 田村 靖

神奈川県秦野市堀山下1番地 株式会社日立製作所汎用コンピュータ事業部内

(73) 発明者 松岡 仁史

神奈川県秦野市堀山下1番地 株式会社日立製作所汎用コンピュータ事業部内

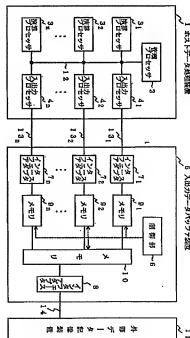
(74) 代理人 弁理士 鈴木 誠

## (54) 【発明の名称】 データ処理システム及び並列コンピュータ

## (57) 【要約】

【目的】 入出力転送速度の違いがホストデータ処理装置と高速な外部データ記憶装置間のデータ転送速度の向上、速度差の吸収を図る。

【構成】 ホストデータ処理装置1は管理プロセッサ2、複数の演算プロセッサ3<sub>1</sub>〜3<sub>n</sub>、複数の入出力プロセッサ4<sub>1</sub>〜4<sub>n</sub>を具備する。入出力データバッファ装置5は制御部6、インタフェースアダプタ7<sub>1</sub>〜7<sub>n</sub>、8、複数のメモリ9<sub>1</sub>〜9<sub>n</sub>及びメモリ10を具備する。管理プロセッサ2は、演算プロセッサからREAD命令あるいはWRITE命令を受け取り、データ入出力動作に必要な複数の入出力プロセッサを選択する。ホストデータ処理装置1と外部データ記憶装置11の間で転送されるデータは複数の分割し、ホストデータ処理装置1の上記選択された複数の入出力プロセッサと入出力データバッファ装置5の複数のメモリ間で並列に転送する。



## 【特許請求の範囲】

【請求項1】 複数の演算プロセッサと複数の入出力プロセッサを持つホストデータ処理装置と、外部データ記憶装置と、複数のバッファメモリを持ち、前記ホストデータ処理装置と前記外部データ記憶装置を相互に接続する入出力データバッファ装置とからなり、前記ホストデータ処理装置の任意演算プロセッサと前記外部データ記憶装置との間で入出力するデータを複数に分割し、前記ホストデータ処理装置の複数の入出力プロセッサと前記入出力データバッファ装置の複数のバッファメモリとの間で並列に転送することを特徴とするデータ処理システム。

【請求項2】 請求項1記載のデータ処理システムにおいて、前記ホストデータ処理装置は、演算プロセッサからのデータ入出力要求に応じ、当該演算プロセッサと前記外部データ記憶装置との間で入出力するデータの分割数及び該分割数分の使用可能な入出力プロセッサを決定する管理プロセッサを有することを特徴とするデータ処理システム。

【請求項3】 各々独立に動作可能な複数の演算プロセッサと複数の入出力プロセッサ、及び、演算プロセッサが外部装置との間で入出力するデータの分割数及び該分割数分の使用可能な入出力プロセッサを決定する管理プロセッサを具備し、任意の演算プロセッサが外部装置との間で入出力するデータを複数に分割し、複数の入出力プロセッサを使用して並列に入出力することを特徴とする並列コンピュータ。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】 本発明は、ホストデータ処理装置と外部データ記憶装置間のデータ転送の速度差を吸収するのに好適なデータ処理システム及び並列コンピュータに関するものである。

## 【0002】

【従来の技術】 従来、ホストコンピュータに代表されるデータ処理装置と、磁気ディスク等の外部データ記憶装置との間は直接接続しており、そのデータ転送速度は両者のうちどちらか遅い方の転送速度に制限されていた。これらの改善策としては、例えば特開平2-93849号公報に記載のように、ホストコンピュータと外部データ記憶装置との間にバッファを設ける方法がある。これにより、ホストコンピュータと外部データ記憶装置の間のデータ転送の速度差の吸収及び外部データ記憶装置の効率的な適用が可能となる。

## 【0003】

【発明が解決しようとする課題】 ところで、上記の従来技術は、ホストコンピュータ側のデータ転送速度が外部データ記憶装置よりも速いという点、及びホストコンピュータ側と外部データ記憶装置との転送パスは一つでシリアル転送を前提としている。

【0004】 一方、最近のデータ処理装置の動向として並列又は超並列コンピュータが注目されている。超並列コンピュータの場合、1プロセッサの処理能力が抑えられることやSCSI等の標準入出力インタフェースの採用等の点で、外部データ記憶装置との入出力転送速度の限界が予想され、高速の外部データ記憶装置が直接接続できないという問題がある。

【0005】 本発明の目的は、ホストデータ処理装置に高速データ転送が可能な外部データ記憶装置を接続する場合、両機器のデータ転送の速度差を吸収し、高速なデータ転送を可能とするデータ処理システムを提供することにある。

【0006】 本発明の他の目的は、複数のプロセッサからなる並列コンピュータにおいて、個々のプロセッサの処理能力が抑えられる場合でも外部装置と高速にデータの入出力を可能とすることにある。

## 【0007】

【課題を解決するための手段】 本発明のデータ処理システムは、複数の演算プロセッサ及び複数の入出力プロセッサを持つホストデータ処理装置と、外部データ記憶装置と、複数のバッファメモリを持ち、ホストデータ処理装置の間に接続される入出力データバッファ装置からなり、ホストデータ処理装置の任意演算プロセッサと外部データ記憶装置との間のデータの入出力を、データ処理装置の複数の入出力プロセッサと入出力データバッファ装置の複数のバッファメモリを使用して並列に行うことを特徴とする。

【0008】 本発明の並列コンピュータは、各々独立に動作可能な複数の演算プロセッサと複数の入出力プロセッサ、及び、演算プロセッサが外部装置との間で入出力するデータの分割数及び該分割数分の使用可能な入出力プロセッサを決定する管理プロセッサを具備し、任意の演算プロセッサが外部装置との間で入出力するデータを複数に分割し、複数の入出力プロセッサを使用して並列に入出力することを特徴とする。

## 【0009】

【作用】 ホストデータ処理装置の或る演算プロセッサが外部データ記憶装置からデータの読み込みを行う場合、外部データ記憶装置から入出力データバッファ装置に転送されたデータは、複数のデータに分割されて、各々、複数のバッファメモリに格納される。これら複数のバッファメモリの各分割されたデータは、ホストデータ処理装置の複数の入出力プロセッサに並列に転送され、当該演算プロセッサに読み込まれる。同様に、ホストデータ処理装置の或る演算プロセッサから外部データ記憶装置へデータの書き込みを行う場合は、該データが複数に分割され、各々、複数の入出力プロセッサから並列に入出力データバッファ装置に転送されて、一旦複数のバッファメモリに格納され、それらのデータが統合されて外部データ記憶装置に転送される。

## 【0010】

【実施例】以下、本発明の一実施例を図面により具体的に説明する。

【0011】図1は、本発明のデータ処理システムの一実施例の全体構成図である。図1において、ホストデータ処理装置1は並列コンピュータで、管理プロセッサ2、複数の演算プロセッサ3<sub>1</sub>〜3<sub>n</sub>、複数の入出力プロセッサ4<sub>1</sub>〜4<sub>n</sub>からなり、各プロセッサはデータ伝送路12で接続され、プロセッサ間で相互にデータ転送が可能となっている。管理プロセッサ2は演算プロセッサ3<sub>1</sub>〜3<sub>n</sub>及び入出力プロセッサ4<sub>1</sub>〜4<sub>n</sub>の動作を管理し、該管理プロセッサ2の制御下で、各演算プロセッサ及び各入出力プロセッサがそれぞれ独立に動作する。なお、管理プロセッサ数と入出力プロセッサ数は同じである必要はない。また、演算プロセッサ3<sub>1</sub>〜3<sub>n</sub>のいずれかが管理プロセッサを兼ねてもよい。

【0012】入出力データバッファ装置5は制御部6、複数のインタフェースアダプタ7<sub>1</sub>〜7<sub>n</sub>、インタフェースアダプタ8、メモリ（バッファメモリ）9<sub>1</sub>〜9<sub>n</sub>、メモリ10からなる。インタフェースアダプタ7<sub>1</sub>〜7<sub>n</sub>、及びメモリ9<sub>1</sub>〜9<sub>n</sub>及びホストデータ処理装置1の入出力プロセッサ4<sub>1</sub>〜4<sub>n</sub>に対応して設けられ、入出力プロセッサ4<sub>1</sub>〜4<sub>n</sub>とインタフェースアダプタ7<sub>1</sub>〜7<sub>n</sub>とはそれぞれデータ伝送路13<sub>1</sub>〜13<sub>n</sub>を介して個別に接続されている。一方、インタフェースアダプタ8はデータ伝送路14を介して外部データ記憶装置11と接続されている。メモリ10は少なくともメモリ9<sub>1</sub>〜9<sub>n</sub>のトータル記憶容量を有し、メモリ9<sub>1</sub>〜9<sub>n</sub>とデータの分配/統合を行う。制御部6は、該入出力データバッファ装置5の全体の制御、メモリ9<sub>1</sub>〜9<sub>n</sub>及びメモリ10の書込み/読み出し動作の制御を行う。

【0013】以下に、図1の構成において、ホストデータ処理装置（並列コンピュータ）1のある演算プロセッサ3<sub>j</sub>が外部データ記憶装置11からデータを読み込む場合の動作（READ動作）、逆にデータ外部記憶装置11へデータを書き込む場合の動作（WRITE動作）を説明する。

【0014】図2は、ホストデータ処理装置1の演算プロセッサ3<sub>j</sub>が外部データ記憶装置11からデータを読み込む場合のデータ転送処理手順を示すシーケンス図である。

【0015】データ処理装置1の演算プロセッサ3<sub>j</sub>が外部データ記憶装置11からデータを読み込む場合、演算プロセッサ3<sub>j</sub>は、管理プロセッサ2に対しREAD命令を発行する。これを受けて管理プロセッサ2は、外部データ記憶装置11との入出力処理で使用する1あるいは複数の入出力プロセッサを選択し、その入出力プロセッサの数及び入出力プロセッサの番号をREAD命令に追加指定する。図3は、READ命令のフォーマットの一例を示す図である。該命令フォーマットは、READ

D命令コマンドヘッダ301、転送元プロセッサ番号302、転送先装置番号303、入出力プロセッサ数304、入出力プロセッサ番号305、READデータ指定506から構成される。このうち、入出力プロセッサ数304及び入出力プロセッサ番号305が、管理プロセッサ2で追加指定される情報である。即ち、管理プロセッサ2では、読み込むデータ量、1つの入出力プロセッサのデータ転送速度、相手外部データ記憶装置11のデータ転送速度などから入出力プロセッサ数304を決定し、この数だけの入出力プロセッサを、入出力プロセッサ4<sub>1</sub>〜4<sub>n</sub>内の使用中でないものから選択して入出力プロセッサ番号305に設定する。ここでは、選択された入出力プロセッサを4<sub>1</sub>〜4<sub>n</sub>とする。転送元プロセッサ番号302はREAD命令を発行した演算プロセッサ3<sub>j</sub>の番号を示す。転送先装置番号303は、READ命令送出先である外部データ記憶装置11に定義された番号である。また、READデータ指定506は、外部データ記憶装置11に読み込むデータを指示するための情報（開始アドレス、転送データ量など）である。

【0016】管理プロセッサ2は、選択した入出力プロセッサ4<sub>1</sub>〜4<sub>n</sub>の一つ（ここでは4<sub>1</sub>とする）に対してREAD命令を転送する。入出力プロセッサ4<sub>1</sub>は、そのREAD命令をデータ伝送路13<sub>1</sub>を介して、入出力データバッファ装置5に転送する。入出力データバッファ装置5は、転送されたREAD命令をそのままインタフェースアダプタ7<sub>1</sub>、メモリ9<sub>1</sub>、メモリ10、インタフェースアダプタ8、データ伝送路14を介して外部データ記憶装置11に転送する。

【0017】READ命令を受けた外部データ記憶装置11は、指定されたデータを読み出し、入出力データバッファ装置5に転送する。この転送データには、READ命令で指定された入出力プロセッサ数及び番号が付加されている。図4は、外部データ記憶装置11から転送されるデータのフォーマットの一例を示す図である。該データフォーマットはデータ転送ヘッダ401、転送先プロセッサ番号402、転送元装置番号403、入出力プロセッサ数404、入出力プロセッサ番号405、データ本体406から構成される。

【0018】外部データ記憶装置11からデータを転送された入出力データバッファ装置5は、一旦、該データをメモリ10に格納した後、制御部6の制御下で、メモリ10からデータを読み出し、該データに付加された入出力プロセッサ数404及び入出力プロセッサ番号405に従い、入出力プロセッサ4<sub>1</sub>〜4<sub>n</sub>に対応するメモリ9<sub>1</sub>〜9<sub>n</sub>に、該データを分割して格納する。次に、入出力データバッファ装置5は、制御部6の制御下で、メモリ9<sub>1</sub>〜9<sub>n</sub>からデータを並列に読み出し、インタフェースアダプタ7<sub>1</sub>〜7<sub>n</sub>、データ伝送路13<sub>1</sub>〜13<sub>n</sub>を介して、ホストデータ処理装置1の指定された複数の入出力プロセッサ4<sub>1</sub>〜4<sub>n</sub>に並列に転送する。図5は、入出力

データバッファ装置5からホストデータ処理装置1の各入出力プロセッサ4<sub>1</sub>~4<sub>k</sub>に転送されるデータのフォーマットの一例を示す図である。該データフォーマットはデータ転送ヘッダ501、転送先プロセッサ番号502、転送元装置番号503、分割データ番号504、分割データ本体505から構成される。ここで、分割データ番号504は分割されたデータの順序を示す番号であり、これにより分割されたデータの順序性が保証される。

【0019】入出力データバッファ装置5からデータを転送されたホストデータ処理装置1の入出力プロセッサ4<sub>1</sub>~4<sub>k</sub>は、それぞれ該データ(分割データ)を管理プロセッサ2に転送する。管理プロセッサ2は、入出力プロセッサ4<sub>1</sub>~4<sub>k</sub>からそれぞれ分割データを受信した後、その分割データ番号504にもとづいてデータを統合し、READ命令発行元の演算プロセッサ3<sub>1</sub>に転送する。

【0020】図6は、ホストデータ処理装置1の演算プロセッサ3<sub>1</sub>が外部データ記憶装置11へデータを書き込む場合のデータ転送処理手順を示すシーケンス図である。

【0021】ホストデータ処理装置1の演算プロセッサ3<sub>1</sub>が外部データ記憶装置11へデータを書き込む場合、演算プロセッサ3<sub>1</sub>は管理プロセッサ2に対しWRITE命令を発行する。図7は、WRITE命令のフォーマットの一例を示す図である。該命令フォーマットはWRITE命令コマンドヘッダ701、転送元プロセッサ番号702、転送元装置番号703、データ本体704から構成される。

【0022】WRITE命令を受けた管理プロセッサ2は、外部データ記憶装置11にデータを転送するのに使用する1つあるいは複数の入出力プロセッサを選択する。ここでは、入出力プロセッサ4<sub>1</sub>~4<sub>k</sub>を選択するとする。なお、選択の基準はREAD命令の場合と同様である。管理プロセッサ2は、選択した入出力プロセッサの数にデータを分割して、WRITE命令を各入出力プロセッサ4<sub>1</sub>~4<sub>k</sub>に転送する。図8は、管理プロセッサ2から各入出力プロセッサ4<sub>1</sub>~4<sub>k</sub>に発行されるWRITE命令のフォーマットの一例を示す図である。該命令フォーマットはWRITE命令コマンドヘッダ801、転送元プロセッサ番号802、転送先装置番号803、分割データ番号804、分割データ本体805から構成される。各入出力プロセッサ4<sub>1</sub>~4<sub>k</sub>は管理プロセッサ2から受け取ったWRITE命令を、データ伝送路13<sub>1</sub>~13<sub>k</sub>を介して入出力データバッファ装置5に対し並列に転送する。

【0023】入出力データバッファ装置5は、ホストデータ処理装置1の入出力プロセッサ4<sub>1</sub>~4<sub>k</sub>から転送された各分割データをインタフェースアダプタ7<sub>1</sub>~7<sub>k</sub>を介して並列にメモリ9<sub>1</sub>~9<sub>k</sub>に格納する。次に、入出力

データバッファ装置5は、制御部6の制御下で、該メモリ9<sub>1</sub>~9<sub>k</sub>の各分割データを読み出し、WRITE命令の分割データ番号にもとづいて統合してメモリ10に格納後、該メモリ10から読み出し、インタフェースアダプタ8、データ伝送路14を介して、外部データ記憶装置11へ転送を行う。その時のデータのフォーマットは、図7に示したものをいれればよい。

【0024】以上、説明したように、図1の構成のデータ処理システムによれば、ホストデータ処理装置(並列コンピュータ)1は、複数の入出力プロセッサ4<sub>1</sub>~4<sub>k</sub>を用いて、入出力データバッファ装置5との間で並列にデータの入出力を行うので、データ転送の高速化を図ることが可能になる。また、該ホストデータ処理装置1の個々の入出力プロセッサ4<sub>1</sub>~4<sub>k</sub>の転送速度が遅く、これに対してデータ外部記憶装置11の転送速度が速い場合でも、該ホストデータ処理装置1とデータ外部記憶装置11との間に設けた入出力データバッファ装置5により、データ転送の速度差を吸収することができる。

【0025】なお、本発明は図示の実施例に限定されるものではなく、次のように変更して実施することもできる。例えば、外部データ記憶装置11に識別番号を持たせ、入出力データバッファ装置5に外部データ記憶装置11用のメモリ10を複数設けることにより、入出力データバッファ装置5に複数台の外部データ記憶装置11の接続を可能とすることができる。また、ホストデータ処理装置1、入出力データバッファ装置5、外部データ記憶装置11を接続するデータ伝送路は、LAN、WAN等の様々な形態のネットワークを適用することも可能である。

#### 【0026】

【発明の効果】以上、説明したように、本発明のデータ処理システムによれば、ホストデータ処理装置の各入出力プロセッサの転送速度が遅くても、複数の入出力プロセッサに対応する複数のバッファメモリを持つ入出力データバッファ装置を用いることにより、高速の外部データ記憶装置を接続することができる。また、ホストデータ処理装置に用いる並列コンピュータは、複数の入出力プロセッサで外部装置と並列にデータの入出力を行うことにより、データ入出力の高速化を図ることができる。

#### 【図面の簡単な説明】

【図1】本発明の一実施例のデータ処理システムの構成図である。

【図2】図1で、ホストデータ処理装置が外部データ記憶装置からデータを読み込む場合のデータ転送処理シーケンスを示す図である。

【図3】READ命令のフォーマットの一例を示す図である。

【図4】外部データ記憶装置がデータを転送する場合のデータフォーマットの一例を示す図である。

【図5】入出力データバッファ装置がホストデータ処理

装置に分割されたデータを転送する場合のデータフォーマットの一例を示す図である。

【図6】図1で、ホストデータ処理装置が外部データ記憶装置へデータを書き込む場合のデータ転送処理シーケンスを示す図である。

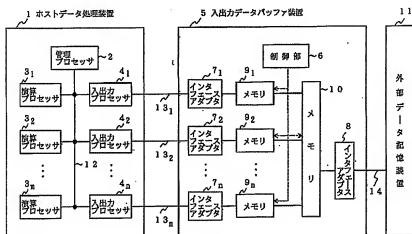
【図7】演算プロセッサが管理プロセッサに転送するWRITE命令のフォーマットの一例を示す図である。

【図8】演算プロセッサが出力プロセッサに転送するWRITE命令のフォーマットの一例を示す図である。

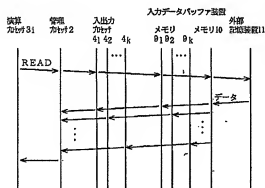
【符号の説明】

- 1 ホストデータ処理装置(並列コンピュータ)
- 2 管理プロセッサ
- 3<sub>1</sub>~3<sub>n</sub> 演算プロセッサ
- 4<sub>1</sub>~4<sub>n</sub> 出力プロセッサ
- 5 入出力データバッファ装置
- 6 制御部
- 7<sub>1</sub>~7<sub>n</sub>、8 インタフェースアダプタ
- 9<sub>1</sub>~9<sub>n</sub> バッファメモリ
- 10 外部データ記憶装置

【図1】



【図2】



【図3】

301	302	303	304	305	306
READ 命令	転送元 プロセッサ 番号	転送先 装置番号	入出力 プロセッサ 番号	入出力 プロセッサ 番号	READ データ指定

【図4】

401	402	403	404	405	406
データ転送 ヘッダ	転送元 プロセッサ 番号	転送先 装置番号	入出力 プロセッサ 番号	入出力 プロセッサ 番号	データ本体

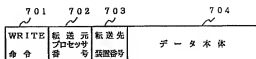
【図5】

501	502	503	504	505
データ転送 ヘッダ	転送元 プロセッサ 番号	転送先 装置番号	分割 データ区	分割データ本体

【図6】



【図7】



【図8】

